Dynkin's Isomorphism with Sign Structure

Kshitij Khare

February 2, 2008

Abstract

The Dynkin isomorphism associates a Gaussian field to a Markov chain. These Gaussian fields can be used as priors for prediction and time series analysis. Dynkin's construction gives rise to Gaussian fields with all non-negative covariances. We extend Dynkin's construction (by introducing a sign structure on the Markov chain) to allow general covariance sign patterns.

1 Introduction

Let $\mathbb{X} = \{X_t\}_{t \geq 0}$ be a reversible Markov process with a countable state space $\mathcal{X}$ and symmetric generator matrix $Q = ((q_{xy}))_{x,y \in \mathcal{X}}$, such that all states are transient. To ensure transience of Q, it is sufficient to assume that either

$$Q \text{ is irreducible and } \sum_{y \in \mathcal{X}} q_{xy} < 0 \text{ for atleast one } x \in \mathcal{X}. \tag{1}$$

or

$$\sum_{y \in \mathcal{X}} q_{xy} < 0 \text{ for every } x \in \mathcal{X}. \tag{2}$$

We provide a proof of this in Section 2. Under both of these assumptions, Q is non-conservative, i.e. from atleast one state there is a positive probability of going to an absorbing "cemetery" Δ (not included in $\mathcal{X}$) and staying there forever. Dynkin [4] associated a Gaussian field $\{Z_x\}_{x \in \mathcal{X}}$ with variance-covariance matrix

$$\Sigma = -Q^{-1}$$

with this Markov process and derived various intersesting properties of this correspondence. Since then, this correspondence has been used in several contexts (See Section 3). All individual covariances of the Gaussian field $\{Z_x\}_{x\in\mathcal{X}}$ are non-negative in this construction. In this paper, we will extend Dynkin's construction to a larger class of variance-covariance matrices, which allow for positive as well as negative covariances.

For this purpose, we introduce a "sign-matrix" $\mathcal{S}$ such that

$$\mathcal{S}(x,x) = 1,\ \mathcal{S}(x,y) = \mathcal{S}(y,x),\ \mathcal{S}(x,y) \in \{-1,1\}\ \forall x \neq y \in \mathcal{X}.$$

If $\mathcal{S}(x,y) = 1$, the transition from x to y is called a positive transition. If $\mathcal{S}(x,y) = -1$, the transition from x to y is called a negative transition. Let S_i denote the random time corresponding to the i^{th} jump for the Markov process $\{X_t\}_{t\geq 0},\ i = 1,2,3,..........$ Define the "sign-process" $\mathbb{H} = \{H_t\}_{t\geq 0}$ by

$$H_t = \prod_{i=1}^{\infty} \mathcal{S}(X_{S_{i-1}}, X_{S_i}) 1_{\{S_i \leq t\}} \text{ (with } S_0 = 0 \text{ and } H_0 = 1).$$

To describe in words, $H_t = 1$ if the number of negative transitions of $\mathbb{X}$ upto time t is even and $H_t = -1$ if the number of negative transitions of $\mathbb{X}$ upto time t is odd. Also, the transition to the "cemetery" Δ from any state is a positive transition by default.

Consider a Gaussian field $\{Z_x^{\mathcal{S}}\}_{x\in\mathcal{X}}$ with variance-covariance matrix

$$\Sigma^{\mathcal{S}} = (-Q \circ \mathcal{S})^{-1}. \tag{3}$$

(We explain what we mean by $(-Q\circ\mathcal{S})^{-1}$ when $\mathcal{X}$ is countably infinite in Section 2). As usual, let

$$l_t^x := \int_0^t 1_{\{X_s = x\}} ds,\ t \geq 0, x \in \mathcal{X}$$

denote the occupation time of the Markov process $\{X_t\}_{t\geq 0}$ in the state x till time t. We prove that for a realization of $\mathbb{X}$ independent of $\{Z_x^{\mathcal{S}}\}_{x\in\mathcal{X}}$ and for each bounded Borel measurable function $F : \mathbb{R}^{|\mathcal{X}|} \to \mathbb{R}$ and $x, y \in \mathcal{X}$,

$$\mathbf{E}\left[Z_x^{\mathcal{S}} Z_y^{\mathcal{S}} F\left(\frac{(Z_u^{\mathcal{S}})^2}{2},\ u \in \mathcal{X}\right)\right] = \int \mathbf{E}\left[F\left(\frac{(Z_u^{\mathcal{S}})^2}{2} + l_\infty^u,\ u \in \mathcal{X}\right) H_\infty\right] d\mu_{xy}, \tag{4}$$

where μ_{xy} is the conditional probability measure given that the process $\{X_t\}_{t\geq 0}$ enters the "cemetery" Δ eventually with y being the last state it stays in before being killed, scaled by a factor of $-Q^{-1}(x,y)$. We also prove identities for conditional prediction of the Gaussian field $\{Z_x^{\mathcal{S}}\}_{x\in\mathcal{X}}$ in terms of the Markov process $\mathbb{X}$. If $A \subset \mathcal{X}$ is finite, then

$$\mathbf{E}[Z_b^{\mathcal{S}} \mid Z_a^{\mathcal{S}},\ a \in A] = \sum_{a\in A} \mathbf{E}_b[1_{\{X_{R_A}=a\}} H_{R_A}] Z_a^{\mathcal{S}},\ \forall b \in \mathcal{X} \setminus A, \tag{5}$$

where R_A is the first time (greater than or equal to S_1) when the Markov process $\mathbb{X}$ hits A, and

$$Cov(Z_b^{\mathcal{S}}, Z_{b'}^{\mathcal{S}} \mid Z_a^{\mathcal{S}},\ a \in A) = \mathbf{E}_b\left[\int_0^\infty 1_{\{X_s=b'\}} H_s 1_{\{s<R_A\}}\right],\ \forall b, b' \in \mathcal{X} \setminus A. \tag{6}$$

Hence, the formulas for Gaussian field predictions in this case can be expressed elegantly in terms of quantities related to the Markov process $\mathbb{X}$.

Note that, Dynkin's construction is a special case of (3) with $\mathcal{S}(x,y) = 1,\ \forall x \neq y \in \mathcal{X}$. In this case $H_t = 1\ \forall t \geq 0$. Also, the version of (4) when $\mathcal{S}(x,y) = 1\ \forall x \neq y \in \mathcal{X}$ is known as the Dynkin's isomorphism theorem. It is remarkable that all changes that arise in the formulas as a result of introducing a "sign-matrix" are reflected by just the "sign-process" $\mathbb{H}$. Note that it is easy to keep track of $\{H_t\}_{t\geq 0}$ while simulating $\{X_t\}_{t\geq 0}$.

2 Preliminaries

We clarify what we mean by inverse of an infinite matrix, atleast the ones that we are dealing with. Let $\tilde{Q}$ be an infinite matrix which can be written as $\tilde{R}(I - \tilde{P})$, where $\tilde{R}$ is a diagonal matrix with negative entries and $|\tilde{P}| := ((|p_{ij}|))_{0\leq i,j<\infty}$ is a sub-Markov matrix satisfying $\sum_{n=0}^\infty |\tilde{P}|^n < \infty$ (i.e. each entry of the matrix is finite). Then the matrix $(\sum_{n=0}^\infty \tilde{P}^n)\tilde{R}^{-1}$ satisfies

$$\tilde{Q}\left(\sum_{n=0}^\infty \tilde{P}^n\right)\tilde{R}^{-1} = \left(\sum_{n=0}^\infty \tilde{P}^n\right)\tilde{R}^{-1}\tilde{Q} = I.$$

Hence, in such cases **we define**

$$\tilde{Q}^{-1} = \left(\sum_{n=0}^\infty \tilde{P}^n\right)\tilde{R}^{-1}.$$

As in the introduction, let $\{X_t\}_{t\geq 0}$ be a reversible Markov process with a countable state space $\mathcal{X}$ with symmetric generator matrix $Q = ((q_{xy}))_{x,y\in\mathcal{X}}$ satisfying (1) or (2). Note that $q_{xy} = q_{yx} \geq 0$ for $x \neq y$, $q_{xx} < 0$ and $\sum_{y\in\mathcal{X}} q_{xy} \leq 0$. Let $\{Y_i\}_{i\geq 0} := \{X_{S_i}\}_{i\geq 0}$ be the embedded discrete-time Markov chain with one step transition probabilities

$$p_{xy} = -\frac{q_{xy}}{q_{xx}} \text{ for } x \neq y,\ p_{xx} = 0.$$

Let $P := ((p_{xy}))_{x,y\in\mathcal{X}}$ and let Q^{diag} denote the diagonal matrix with diagonal entries same as Q. Then

$$Q = Q^{diag}(I - P). \tag{7}$$

We describe a typical path of $\{X_t\}_{t\geq 0}$. The process starts at an initial state Y_0. The process stays at Y_n during $[S_n, S_{n+1})$ for $n = 0, 1, 2, \ldots\ldots\ldots.$ and at a random time $\xi = S_{\eta+1}$ jumps from the state Y_η to the "cemetery" Δ (not included in $\mathcal{X}$), and stays there forever. (Here ξ takes non-negative real values and η takes non-negative integer values). The value of ξ (and hence η) can be infinite for certain sample paths, in which case the path does not terminate. Note that the difference $1 - \sum_{y\in\mathcal{X}} p_{xy}$ represents the probability $p(x, \Delta)$ of a jump from x to the "cemetery" Δ for the embedded Markov chain $\{Y_m\}_{m\geq 0}$. Also, conditional on $\{Y_m\}_{m\geq 0}$, the intermediate jump times $\{S_{i+1} - S_i\}_{i\geq 0}$ are independent and have $Exponential(-q_{Y_iY_i})$ distribution for $i = 0, 1, 2, \ldots\ldots\ldots.$

Let us prove by the method of contradiction that under (1) or (2), $\sum_{n=0}^{\infty} P^n < \infty$ and hence Q^{-1} exists. For this it is enough to show that all states are transient. Suppose $x \in \mathcal{X}$ is recurrent. The assumptions (1) or (2) imply that $\exists n$ such that $P^n(x, \Delta) > 0$. This implies by [3, Theorem 3.4] that starting from the "cemetery" Δ there is a positive probability of reaching the state x, which is a contradiction as the "cemetery" Δ is absorbing. Hence, $\sum_{n=0}^{\infty} P^n < \infty$ and $Q^{-1} = (\sum_{n=0}^{\infty} P^n)(Q^{diag})^{-1}$ exists.

The lemma below relates Q^{-1} to the expected infinite occupation times for the process $\{X_t\}_{t\geq 0}$.

Lemma 2.1

$$-Q^{-1}(x, y) = \mathbf{E}_x[l^y_\infty],\ \forall x, y \in \mathcal{X}.$$

Proof Firstly, $Q = Q^{diag}(I - P)$ leads to

$$Q^{-1} = \left(\sum_{n=0}^{\infty} P^n \right) (Q^{diag})^{-1}.$$

By decomposing the path of the Markov chain in terms of jump times, we get,

$$\begin{aligned}
\mathbf{E}_x \left[\int_0^{\infty} 1_{\{X_s = y\}} ds \right] &= \mathbf{E}_x \left[\sum_{i=0}^{\infty} (S_{i+1} - S_i) 1_{\{X_{S_i} = y\}} \right] \\
&= \mathbf{E}_x \left[\sum_{i=0}^{\infty} (S_{i+1} - S_i) 1_{\{Y_i = y\}} \right] \\
&= \mathbf{E}_x \left[\sum_{i=0}^{\infty} \mathbf{E}_x \left[(S_{i+1} - S_i) \mid \{Y_m\}_{m \geq 0} \right] 1_{\{Y_i = y\}} \right] \\
&= \mathbf{E}_x \left[\sum_{i=0}^{\infty} \frac{-1}{q_{Y_i Y_i}} 1_{\{Y_i = y\}} \right]
\end{aligned}$$

The previous equality follows from the fact that conditional on $\{Y_m\}_{m \geq 0}$, the intermediate jump times $\{S_{i+1} - S_i\}_{i \geq 0}$ are independent and have $Exponential(-q_{Y_i Y_i})$ distribution for $i = 0, 1, 2,$ ....
This gives

$$\begin{aligned}
\mathbf{E}_x \left[\int_0^{\infty} 1_{\{X_s = y\}} ds \right] &= -\frac{1}{q_{yy}} \mathbf{E}_x \left[\sum_{i=0}^{\infty} 1_{\{Y_i = y\}} \right] \\
&= -\frac{1}{q_{yy}} \sum_{i=0}^{\infty} \mathbf{P}_x \{Y_i = y\} \\
&= -\frac{1}{q_{yy}} \sum_{i=0}^{\infty} P^n(x, y)
\end{aligned}$$

Note that $\mathbf{P}_x$ denotes the probability distribution starting at the state x, while P is the transition matrix for the embedded Markov chain $\{Y_m\}_{m \geq 0}$. Hence,

$$\begin{aligned}
\mathbf{E}_x[l_{\infty}^y] &= \mathbf{E}_x \left[\int_0^{\infty} 1_{\{X_s = y\}} ds \right] \\
&= -(Q^{diag}(I - P))^{-1}(x, y) \\
&= -Q^{-1}(x, y) \blacksquare
\end{aligned}$$

3 History

In a series of papers Dynkin [4, 5, 6, 7] proposed and built on his construction as a connection between random fields and Markov processes. Marcus and Rosen have used the refined knowledge about Gaussian fields (eg. continuity of sample paths) to develop fine properties of symmetric Markov processes (eg. continuity of local times) using Dynkin's construction. Their book [11] gives a detailed and accessible account of their methods. Sheppard [13] uses Dynkin's isomorphism to give a proof of the Ray-Knight theorem on the Markovianity of one-dimensional diffusions. The properties of Markov processes can be utilized for analyzing the corresponding Gaussian fields. Ylvisaker [14] uses Gaussian fields (amenable to Dynkin's isomorphism) as Bayesian priors for prediction and design problems, and makes use of the formulas relating the prediction properties of the Gaussian field to the corresponding Markov process. Bolthausen [1] uses Dynkin's isomorphism as a tool in analyzing the limiting behaviour of the Gaussian free field. Eisenbaum [9] and also Marcus and Rosen [11] have established variants of Dynkin's isomrphism. In the case of diffusions, Eisenbaum [8] shows that Dynkin's isomorphism theorem and the Ray-Knight theorems can be derived from each other. In [10] the authors use an unconditional version of Dynkin's isomorphism to obtain a Ray-Knight theorem for a class of symmetric Markov processes. Diaconis and Evans [2] proposed a different construction by looking at $-Q$ as the variance-covariance matrix instead of $-Q^{-1}$. Their construction yields Gaussian fields with negative individual covariances.

4 Generalization of Dynkin's Isomorphism

We again consider a Markov process $\{X_t\}_{t\geq 0}$ with a countable state space $\mathcal{X}$ and with a generator matrix Q as in Section 2. We introduce a "sign-matrix" $\mathcal{S}$ such that

$$\mathcal{S}(x,x) = 1,\ \mathcal{S}(x,y) = \mathcal{S}(y,x),\ \mathcal{S}(x,y) \in \{-1,1\}\ \forall x \neq y \in \mathcal{X}.$$

As explained in the introduction, if $\mathcal{S}(x,y) = 1$, the transition from x to y is called a **positive transition**. If $\mathcal{S}(x,y) = -1$, the transition from x to y is called a **negative transition**. The "sign-process" $\mathbb{H} = \{H_t\}_{t\geq 0}$ is defined

by

$$H_t = \prod_{i=1}^{\infty} \mathcal{S}(Y_{i-1}, Y_i) 1_{\{S_i \le t\}},\ \text{(with } H_0 = 1),$$

and $H_t = 1$ if the number of negative transitions of $\mathbb{X}$ upto time t is even and $H_t = -1$ if the number of negative transitions of $\mathbb{X}$ upto time t is odd. Also, the transition to the "cemetery" Δ from any state is a positive transition by default.

Example If $\mathcal{S}(x,y) = 1\ \forall x \neq y \in \mathcal{X}$, then $H_t \equiv 1\ \forall t \geq 0$ and $\Sigma = -Q^{-1}$.

Example If $\mathcal{S}(x,y) = -1\ \forall x \neq y \in \mathcal{X}$, then $H_t = (-1)^{|\{i \geq 1 \colon S_i \leq t\}|}\ \forall t \geq 0$ and $\Sigma = -(I+P)^{-1}(Q^{diag})^{-1}$ ($\because$ From (7)). For this particular case, the hyper-process $\{H_t\}_{t \geq 0}$ is 1 between $[S_{2i}, S_{2i+1})$ and -1 between $[S_{2i+1}, S_{2i+2})$ for $i \geq 0$.

Define

$$\tilde{l}_t^x := \int_0^t 1_{\{X_s = x, H_s = 1\}} ds - \int_0^t 1_{\{X_s = x, H_s = -1\}} ds = \int_0^t 1_{\{X_s = 1\}} H_s ds,\ \ t \geq 0,\ x \in \mathcal{X}.$$

We interpret $\tilde{l}_t^x$ as the **net occupation time** in the state x till time t (with $\int_0^t 1_{\{X_s = x, H_s = 1\}} ds$ and $\int_0^t 1_{\{X_s = x, H_s = -1\}} ds$ interpreted as the negative and positive occupation times respectively).

Define the matrix

$$\Sigma^{\mathcal{S}} := (-Q \circ \mathcal{S})^{-1}.$$

Here $\circ$ denotes Hadamard product i.e. elementwise product of the two matrices. Note that $|P \circ \mathcal{S}| = P$ and $\sum_{n=0}^{\infty} P^n < \infty$. Hence $(-Q \circ \mathcal{S})^{-1}$ exists. Also, since $-Q \circ \mathcal{S}$ is a diagonally dominant matrix with positive diagonal entries, hence $\Sigma^{\mathcal{S}}$ is positive definite. Note that, an infinite matrix is defined to be positive definite if all its finite principal submatrices are positive definite.

Lemma 4.1

$$\Sigma^{\mathcal{S}}(x,y) = \mathbf{E}_x \left[\int_0^{\infty} 1_{\{X_s = y\}} H_s ds \right] = \mathbf{E}_x [\tilde{l}_{\infty}^y],\ \forall x, y \in \mathcal{X}.$$

i.e. $\Sigma^{\mathcal{S}}(x,y)$ *is the expected net occupation time at* y *starting at* x.

Proof Firstly,

$$\begin{aligned} Q \circ \mathcal{S} &= (Q^{diag}(I-P)) \circ \mathcal{S} \\ &= Q^{diag}(I - P \circ \mathcal{S}) \end{aligned}$$

This gives

$$\Sigma^{\mathcal{S}} = (-Q \circ \mathcal{S})^{-1} = (P \circ \mathcal{S} - I)^{-1}(Q^{diag})^{-1}.$$

By decomposing the path of the Markov chain in terms of jump times, we get,

$$\begin{aligned} \mathbf{E}_x\left[\int_0^\infty 1_{\{X_s=y\}} H_s ds\right] &= \mathbf{E}_x\left[\sum_{i=0}^\infty (S_{i+1}-S_i) 1_{\{X_{S_i}=y\}} H_{S_i}\right] \\ &= \mathbf{E}_x\left[\sum_{i=0}^\infty (S_{i+1}-S_i) 1_{\{Y_i=y\}} H_{S_i}\right] \\ &= \mathbf{E}_x\left[\sum_{i=0}^\infty \mathbf{E}_x\left[(S_{i+1}-S_i) \mid \{Y_m\}_{m\geq 0}\right] 1_{\{Y_i=y\}} H_{S_i}\right] \end{aligned}$$

The previous equality follows from the fact that

$$H_{S_i} = \prod_{j=1}^{i} \mathcal{S}(Y_{j-1}, Y_j)$$

is a function of $\{Y_m\}_{m\geq 0}$. This gives,

$$\begin{aligned} \mathbf{E}_x\left[\int_0^\infty 1_{\{X_s=y\}} H_s ds\right] &= \mathbf{E}_x\left[\sum_{i=0}^\infty \frac{-1}{q_{Y_i Y_i}} 1_{\{Y_i=y\}} H_{S_i}\right] \\ &= -\frac{1}{q_{yy}} \mathbf{E}_x\left[\sum_{i=0}^\infty 1_{\{Y_i=y\}} H_{S_i}\right] \\ &= -\frac{1}{q_{yy}} \sum_{i=0}^\infty \mathbf{E}_x\left[1_{\{Y_i=y\}} H_{S_i}\right] \end{aligned}$$

The exchange of sum and expectation is justified by the fact that $|H_s| = 1$ and $\sum_{i=0}^\infty \mathbf{E}_x\left[1_{\{Y_i=y\}}\right] = (I-P)^{-1}(x,y) < \infty$.

Let us calculate $\mathbf{E}_x\left[1_{\{Y_i=y\}}H_{S_i}\right]$.

$$\begin{aligned}
\mathbf{E}_x\left[1_{\{Y_i=y\}}H_{S_i}\right] &= \mathbf{E}_x\left[1_{\{Y_i=y\}}\prod_{j=1}^{i}\mathcal{S}(Y_{j-1},Y_j)\right] \\
&= \sum_{y_1,y_2,\ldots,y_{i-1}\in\mathcal{X}}\prod_{j=1}^{i}p_{y_{j-1}y_j}\prod_{j=1}^{i}\mathcal{S}(y_{j-1},y_j) \text{ where } y_0=x, y_i=y. \\
&= \sum_{y_1,y_2,\ldots,y_{i-1}\in\mathcal{X}}\prod_{j=1}^{i}p_{y_{j-1}y_j}\mathcal{S}(y_{j-1},y_j) \text{ where } y_0=x, y_i=y. \\
&= (P\circ\mathcal{S})^i(x,y)
\end{aligned}$$

Hence,

$$\begin{aligned}
\mathbf{E}_x[\tilde{l}^y_\infty] &= \mathbf{E}_x[\int_0^\infty 1_{\{X_s=y\}}H_s ds] \\
&= -\frac{1}{q_{yy}}\sum_{i=0}^{\infty}(P\circ\mathcal{S})^i(x,y) \\
&= -\frac{1}{q_{yy}}(I-P\circ\mathcal{S})^{-1}(x,y) \\
&= -(Q^{diag}(I-P\circ\mathcal{S}))^{-1}(x,y) \\
&= \Sigma^{\mathcal{S}}(x,y)
\end{aligned}$$

The proof is complete. ■

We next prove the isomorphism theorem (4) for a zero mean Gaussian process $\{Z^{\mathcal{S}}_x\}_{x\in\mathcal{X}}$ with variance-covriance matrix $\Sigma=(-Q\circ\mathcal{S})^{-1}$ and an independent realization $\{X_t\}_{t\geq 0}$ of the Markov process with generator Q.

4.1 The finite case

We consider the case when $\mathcal{X}$ is finite. We proceed similarily as Dynkin [4] and first consider functions of the form

$$F_{\underline{d}}(\underline{w}) = e^{-\sum_{u\in\mathcal{X}} d_u w_u},$$

where $\underline{d}=\{d_u\}_{u\in\mathcal{X}}$ is arbitrary with $d_u\geq 0,\ \forall u\in\mathcal{X}$. Let D denote the diagonal matrix with diagonal entries $\{d_u\}_{u\in\mathcal{X}}$. Then,

$$\mathbf{E}\left[Z_xZ_y e^{-\sum_{u\in\mathcal{X}} d_u\frac{Z_u^2}{2}}\right] = \int_{\mathbb{R}^{|\mathcal{X}|}}\frac{z_xz_y}{(\sqrt{2\pi})^{|\mathcal{X}|}det(-Q\circ\mathcal{S})}e^{-\frac{\underline{z}^TD\underline{z}-\underline{z}^T(Q\circ\mathcal{S})\underline{z}}{2}}d\underline{z}$$

$$
\begin{aligned}
&= \frac{det(D-Q\circ\mathcal{S})}{det(-Q\circ\mathcal{S})}\int_{\mathbb{R}^{|\mathcal{X}|}} \frac{z_x z_y}{(\sqrt{2\pi})^{|\mathcal{X}|}det(D-Q\circ\mathcal{S})} e^{-\frac{\underline{z}^T(D-Q\circ\mathcal{S})\underline{z}}{2}} d\underline{z} \\
&= \frac{det((D-Q)\circ\mathcal{S})}{det(-Q\circ\mathcal{S})}\int_{\mathbb{R}^{|\mathcal{X}|}} \frac{z_x z_y}{(\sqrt{2\pi})^{|\mathcal{X}|}det((D-Q)\circ\mathcal{S})} e^{-\frac{\underline{z}^T((D-Q)\circ\mathcal{S})\underline{z}}{2}} d\underline{z} \\
&= \frac{det((D-Q)\circ\mathcal{S})}{det(-Q\circ\mathcal{S})}((D-Q)\circ\mathcal{S})^{-1}(x,y)
\end{aligned}
$$

Note that $\mathbf{E}\left[e^{-\sum_{u\in\mathcal{X}} d_u \frac{Z_u^2}{2}}\right] = \frac{det((D-Q)\circ\mathcal{S})}{det(-Q\circ\mathcal{S})}$ by a similiar calculation as above.

Hence,

$$
\mathbf{E}\left[Z_x Z_y e^{-\sum_{u\in\mathcal{X}} d_u \frac{Z_u^2}{2}}\right] = \mathbf{E}\left[e^{-\sum_{u\in\mathcal{X}} d_u \frac{Z_u^2}{2}}\right]((D-Q)\circ\mathcal{S})^{-1}(x,y) \tag{8}
$$

Note that $-(D-Q)$ is the generator of a Markov process $\{\bar{X}_t\}_{t\geq 0}$ with the same structure as $\{X_t\}_{t\geq 0}$ except that at every state $x \in \mathcal{X}$, there is an additional killing rate of d_x. Let $\{\bar{Y}_m\}_{m\geq 0}$ be the embedded discrete-time Markov chain and $\{\bar{H}_t\}_{t\geq 0}$ the hyper-process corresponding to $(\{\bar{X}_t\}_{t\geq 0}, \mathcal{S})$. Let us establish the change of measure formula from $\{\bar{Y}_m\}_{m\geq 0}$ to $\{Y_m\}_{m\geq 0}$.

$$
\begin{aligned}
P_x\{\bar{Y}_1 = y_1, \bar{Y}_2 = y_2, ..., \bar{Y}_n = y_n\} &= \prod_{i=1}^{n} \frac{q_{y_{i-1}y_i}}{-q_{y_{i-1}y_{i-1}} + d_{y_{i-1}}} \text{ (with } y_0 = x) \\
&= \left(\prod_{i=1}^{n} \frac{-q_{y_{i-1}y_{i-1}}}{-q_{y_{i-1}y_{i-1}} + d_{y_{i-1}}}\right) P_x\{Y_1 = y_1, Y_2 = y_2, ..., Y_n = y_n\}
\end{aligned}
$$

Hence,

$$
\mathbf{E}_x[F(\bar{Y}_1, \bar{Y}_2, ..., \bar{Y}_n)] = \mathbf{E}_x\left[\left(\prod_{i=1}^{n} \frac{-q_{y_{i-1}y_{i-1}}}{-q_{y_{i-1}y_{i-1}} + d_{y_{i-1}}}\right) F(Y_1, Y_2, ..., Y_n)\right] \tag{9}
$$

for each bounded Borel measurable function F.

As in the introduction, let us define the measure μ_{xy} by

$$
\mu_{xy}\{C\} = -Q^{-1}(x,y)\mathbf{P}_x\{C \mid \eta < \infty, Y_\eta = y\}. \tag{10}
$$

It is the appropriately scaled conditional probability measure given that the process $\mathbb{X}$ enters the "cemetery" Δ eventually with y being the last state it

stays in before being killed. Note that,

$$\begin{aligned}
& \mu_{xy}\{Y_1 = y_1, Y_2 = y_2, Y_3 = y_3, ..., Y_n = y_n, \eta = n\} \\
= \quad & -Q^{-1}(x,y)\frac{P_x\{Y_1 = y_1, Y_2 = y_2, Y_3 = y_3, ..., Y_n = y_n, Y_{n+1} = \Delta\}1_{\{y_n = y\}}}{P_x\{\eta < \infty, Y_\eta = y\}} \\
= \quad & -Q^{-1}(x,y)\frac{P_x\{Y_1 = y_1, Y_2 = y_2, ..., Y_n = y_n\}1_{\{y_n=y\}}p(y,\Delta)}{\sum_{n=0}^{\infty} P_x\{Y_n = y, Y_{n+1} = \Delta\}} \\
= \quad & \frac{-Q^{-1}(x,y)p(y,\Delta)}{\sum_{n=0}^{\infty} P_x\{Y_n = y\}p(y,\Delta)} P_x\{Y_1 = y_1, Y_2 = y_2, ..., Y_n = y_n\}1_{\{y_n=y\}} \\
= \quad & \frac{1}{-q_{yy}} P_x\{Y_1 = y_1, Y_2 = y_2, ..., Y_n = y_n\}1_{\{y_n=y\}}
\end{aligned}$$

The previous equality follows from the fact that $-Q^{-1}(x,y) = \frac{1}{-q_{yy}} \sum_{n=0}^{\infty} P_x\{Y_n = y\}$.
Hence,

$$\int F(Y_1, Y_2, ..., Y_\eta)1_{\{\eta=n\}}d\mu_{xy} = \frac{1}{-q_{yy}}\mathbf{E}_x\left[F(Y_1, Y_2, ..., Y_n)1_{\{Y_n=y\}}\right] \qquad (11)$$

for each bounded Borel measurable function F.

We combine these results to evaluate $((D-Q)\circ\mathcal{S})^{-1}(x,y)$ in terms of $\{l_\infty^u\}_{u\in\mathcal{X}}$ and H_∞. Let μ_{xy} be the measure defined in (10).

Lemma 4.2

$$((D-Q)\circ\mathcal{S})^{-1}(x,y) = \int e^{-\sum_{u\in\mathcal{X}} d_u l_\infty^u} H_\infty d\mu_{xy}.$$

Proof Firstly, we observe that under μ_{xy}, $\eta < \infty$ and hence $H_\infty = H_{S_\eta}$, which is a measurable function of $Y_1, Y_2, ..., Y_\eta$. Also,

$$\sum_{u\in\mathcal{X}} d_u l_\infty^u = \sum_{i=0}^{\eta} d_{Y_i}(S_{i+1} - S_i).$$

Hence,

$$\int e^{-\sum_{u\in\mathcal{X}} d_u l_\infty^u} H_\infty d\mu_{xy}$$

$$
\begin{aligned}
&= \sum_{n=0}^{\infty} \int e^{-\sum_{i=0}^{n} d_{Y_i}(S_{i+1}-S_i)} H_{S_n} 1_{\{\eta=n\}} d\mu_{xy} \\
&= \sum_{n=0}^{\infty} \frac{1}{-q_{yy}} \mathbf{E}_x \left[e^{-\sum_{i=0}^{n} d_{Y_i}(S_{i+1}-S_i)} H_{S_n} 1_{\{Y_n=y\}} \right] \ (\because \text{From } (11)) \\
&= \sum_{n=0}^{\infty} \frac{1}{-q_{yy}} \mathbf{E}_x \left[\mathbf{E}_x \left[e^{-\sum_{i=0}^{n} d_{Y_i}(S_{i+1}-S_i)} \mid \{Y_m\}_{m\geq 0} \right] H_{S_n} 1_{\{Y_n=y\}} \right] \\
&= \sum_{n=0}^{\infty} \frac{1}{-q_{yy}} \mathbf{E}_x \left[\left(\prod_{i=0}^{n} \frac{-q_{Y_iY_i}}{-q_{Y_iY_i}+d_{Y_i}} \right) H_{S_n} 1_{\{Y_n=y\}} \right]
\end{aligned}
$$

The previous equality follows from the fact that conditioned on $\{Y_m\}_{m\geq 0}$, the intermediate jump times $\{S_{i+1}-S_i\}_{i\geq 0}$ are independent and have $Exponential(-q_{Y_iY_i})$ distribution for $i = 0, 1, 2, ...$
Hence, with $\{\bar{S}_i\}_{i\geq 0}$ denoting the random transition times for $\{\bar{X}_t\}_{t\geq 0}$, we get,

$$
\begin{aligned}
\int e^{-\sum_{u\in\mathcal{X}} d_u l_\infty^u} H_\infty d\mu_{xy} &= \sum_{n=0}^{\infty} \frac{1}{-q_{yy}+d_y} \mathbf{E}_x \left[\left(\prod_{i=1}^{n} \frac{-q_{Y_{i-1}Y_{i-1}}}{-q_{Y_{i-1}Y_{i-1}}+d_{Y_{i-1}}} \right) H_{S_n} 1_{\{Y_n=y\}} \right] \\
&= \sum_{n=0}^{\infty} \frac{1}{-q_{yy}+d_y} \mathbf{E}_x \left[\bar{H}_{\bar{S}_n} 1_{\{\bar{Y}_n=y\}} \right] \ (\because \text{From } (9)) \\
&= \mathbf{E}_x \left[\sum_{n=0}^{\infty} \frac{1_{\{\bar{Y}_n=y\}} \bar{H}_{\bar{S}_n}}{-q_{yy}+d_y} \right] \\
&= \mathbf{E}_x \left[\int_0^\infty 1_{\{\bar{X}_s=y\}} \bar{H}_s ds \right] \\
&= ((D-Q)\circ\mathcal{S})^{-1}(x,y)
\end{aligned}
$$

The previous equality follows from the fact that $\{\bar{X}_t\}_{t\geq 0}$ is a Markov process with generator $-(D-Q)$ which satisfies (1) or (2).
Hence proved. ■

It follows from this claim and (8) that

$$
\mathbf{E}\left[Z_x^{\mathcal{S}} Z_y^{\mathcal{S}} e^{-\sum_{u\in\mathcal{X}} d_u \frac{(Z_u^{\mathcal{S}})^2}{2}} \right] = \mathbf{E}\left[e^{-\sum_{u\in\mathcal{X}} d_u \frac{(Z_u^{\mathcal{S}})^2}{2}} \right] \int e^{-\sum_{u\in\mathcal{X}} d_u l_\infty^u} H_\infty d\mu_{xy}.
$$

Hence,

$$\mathbf{E}\left[Z_x^{\mathcal{S}} Z_y^{\mathcal{S}} e^{-\sum_{u\in\mathcal{X}} d_u \frac{(Z_u^{\mathcal{S}})^2}{2}}\right] = \int \mathbf{E}\left[e^{-\sum_{u\in\mathcal{X}} d_u \left(\frac{(Z_u^{\mathcal{S}})^2}{2} + l_\infty^u\right)} H_\infty\right] d\mu_{xy}.$$

The set of functions $F_{\underline{d}}(\underline{w}) = e^{-\sum_{u\in\mathcal{X}} d_u w_u}$, where $\underline{d} = \{d_u\}_{u\in\mathcal{X}}$ is arbitrary with $d_u \geq 0,\ \forall u \in \mathcal{X}$, generate the Borel σ-algebra in $\mathbb{R}^{|\mathcal{X}|}$ and they form a closed class under multiplication. Also, the set of functions F for which

$$\mathbf{E}\left[Z_x^{\mathcal{S}} Z_y^{\mathcal{S}} F\left(\frac{(Z_u^{\mathcal{S}})^2}{2},\ u \in \mathcal{X}\right)\right] = \int \mathbf{E}\left[F\left(\frac{(Z_u^{\mathcal{S}})^2}{2} + l_\infty^u,\ u \in \mathcal{X}\right) H_\infty\right] d\mu_{xy},$$

is a linear space closed under bounded convergence and under monotone convergence. Hence, for each bounded Borel measurable function $F : \mathbb{R}^{|\mathcal{X}|} \to \mathbb{R}$,

$$\mathbf{E}\left[Z_x^{\mathcal{S}} Z_y^{\mathcal{S}} F\left(\frac{(Z_u^{\mathcal{S}})^2}{2},\ u \in \mathcal{X}\right)\right] = \int \mathbf{E}\left[F\left(\frac{(Z_u^{\mathcal{S}})^2}{2} + l_\infty^u,\ u \in \mathcal{X}\right) H_\infty\right] d\mu_{xy}, \tag{12}$$

where $\{Z_x^{\mathcal{S}}\}_{x\in\mathcal{X}}$ is a zero mean Gaussian field with variance-covariance matrix given by $(-Q \circ \mathcal{S})^{-1}$, and $\{l_\infty^u\}_{u\in\mathcal{X}}$ are the occupation times of a realization $\{X_t\}_{t\geq 0}$ independent of $\{Z_x^{\mathcal{S}}\}_{x\in\mathcal{X}}$ of a Markov process with generator Q. Also, in this case the map $(Q, \mathcal{S}) \to (-Q \circ \mathcal{S})^{-1}$ is one-to-one, because

$$\begin{aligned}(-Q_1 \circ \mathcal{S}_1)^{-1} = (-Q_2 \circ \mathcal{S}_2)^{-1} \quad &\Leftrightarrow \quad Q_1 \circ \mathcal{S}_1 = Q_2 \circ \mathcal{S}_2 \\ &\Leftrightarrow \quad Q_1 = Q_2,\ \mathcal{S}_1 = \mathcal{S}_2\end{aligned}$$

The previous statement is justified by the fact that $Q_i,\ i = 1,2$ have negative off-diagonal entries and $\mathcal{S}_i,\ i = 1,2$ have entries that equal 1 or -1.

Let us now turn our attention to the problem of predicting the above Gaussian field given observations in a proper subset $A \subset \mathcal{X}$, and proving the identities (5) and (6). We do not require the assumption of independence of $\{X_t\}_{t\geq 0}$ and $\{Z_x^{\mathcal{S}}\}_{x\in\mathcal{X}}$ for these calculations. Let $B := \mathcal{X} \setminus A$. Note that,

$$\mathbf{E}[Z_B^{\mathcal{S}} \mid Z_A^{\mathcal{S}}] = \Sigma_{BA} {\Sigma_{AA}}^{-1} Z_A^{\mathcal{S}}.$$

Since $\Sigma^{\mathcal{S}} = (-Q \circ \mathcal{S})^{-1}$, it follows that

$$\Sigma_{BA}^{\mathcal{S}} {\Sigma_{AA}^{\mathcal{S}}}^{-1} = -\{(-Q \circ \mathcal{S})_{BB}\}^{-1} (-Q \circ \mathcal{S})_{BA}.$$

By slightly detailed, but straightforward matrix computations as in Lemma 4.1, it follows that

$$\{(-Q \circ \mathcal{S})_{BB}\}^{-1}(b, b') = \mathbf{E}_b \left[\sum_{i=0}^{\infty} \frac{1_{\{S_i < R_A\}} 1_{\{Y_i = b'\}} H_{S_i}}{-q_{b'b'}} \right].$$

where R_A is the first time (greater than equal to S_1) that the Markov chain $\{X_t\}_{t \geq 0}$ hits A and H_{S_i} is the "sign-process" evaluated at the i^{th} jump time S_i, for $i \geq 0$. Hence,

$$\begin{aligned}
\Sigma_{BA}^{\mathcal{S}} {\Sigma_{AA}^{\mathcal{S}}}^{-1}(b, a) &= \sum_{b' \in B} \mathbf{E}_b \left[\sum_{i=0}^{\infty} 1_{\{S_i < R_A\}} 1_{\{Y_i = b'\}} H_{S_i} \right] \left(\frac{q_{b'a}}{-q_{b'b'}} \right) \mathcal{S}(b', a) \\
&= \sum_{i=0}^{\infty} \sum_{b' \in B} \mathbf{E}_b \left[1_{\{S_i < R_A\}} 1_{\{Y_i = b'\}} 1_{\{Y_{i+1} = a\}} H_{S_{i+1}} \right]
\end{aligned}$$

The previous equality follows by conditioning, the Markov property and $P_{b'}\{Y_1 = a\} = \frac{q_{b'a}}{-q_{b'b'}} = p(b', a)$.
Hence,

$$\begin{aligned}
\Sigma_{BA}^{\mathcal{S}} {\Sigma_{AA}^{\mathcal{S}}}^{-1}(b, a) &= \mathbf{E}_b \left[\sum_{i=0}^{\infty} 1_{\{S_i < R_A, Y_i \in B, Y_{i+1} = a\}} H_{S_{i+1}} \right] \\
&= \mathbf{E}_b \left[\sum_{i=0}^{\infty} 1_{\{R_A = S_{i+1}, Y_{i+1} = a\}} H_{S_{i+1}} \right]
\end{aligned}$$

It follows that,

$$\Sigma_{BA}^{\mathcal{S}} {\Sigma_{AA}^{\mathcal{S}}}^{-1}(b, a) = \{(-Q \circ \mathcal{S})_{BB}\}^{-1} (-Q \circ \mathcal{S})_{BA} = \mathbf{E}_b \left[1_{\{X_{R_A} = a\}} H_{R_A} \right]. \quad (13)$$

Hence, $\forall b \in B$,

$$\mathbf{E}[Z_b^{\mathcal{S}} \mid Z_a^{\mathcal{S}},\ a \in A] = \sum_{a \in A} \mathbf{E}_b \left[1_{\{X_{R_A} = a\}} H_{R_A} \right] Z_a^{\mathcal{S}}. \quad (14)$$

Also,

$$\begin{aligned}
Var[Z_B^{\mathcal{S}} \mid Z_A^{\mathcal{S}}] &= \Sigma_{BB}^{\mathcal{S}} - \Sigma_{BA}^{\mathcal{S}} {\Sigma_{AA}^{\mathcal{S}}}^{-1} \Sigma_{AB}^{\mathcal{S}} \\
&= \{(-Q \circ \mathcal{S})_{BB}\}^{-1}
\end{aligned}$$

Hence,

$$Cov(Z_b^{\mathcal{S}}, Z_{b'}^{\mathcal{S}} \mid Z_a^{\mathcal{S}},\ a \in A) = \mathbf{E}_b \left[\sum_{i=0}^{\infty} \frac{1_{\{S_i < R_A\}} 1_{\{Y_i = b'\}} H_{S_i}}{-q_{b'b'}} \right],\ \forall b, b' \in B.$$

It follows that,

$$Cov(Z_b^{\mathcal{S}}, Z_{b'}^{\mathcal{S}} \mid Z_a^{\mathcal{S}},\ a \in A) = \mathbf{E}_b \left[\int_0^{\infty} 1_{\{X_s = b'\}} H_s 1_{\{s < R_A\}} ds \right],\ \forall b, b' \in B. \tag{15}$$

4.2 The infinite case

We now deal with the case when $\mathcal{X}$ is countably infinite. To prove (4) in this case, arbitrarily fix a finite subset $\mathcal{X}_f \subset \mathcal{X}$. Note that the variance-covariance matrix for $\{Z_x^{\mathcal{S}}\}_{x \in \mathcal{X}_f}$ is given by

$$\Sigma_f := (-(Q \circ \mathcal{S})_{\mathcal{X}_f \mathcal{X}_f} + (Q \circ \mathcal{S})_{\mathcal{X}_f \mathcal{X}_f^c} \{(Q \circ \mathcal{S})_{\mathcal{X}_f^c \mathcal{X}_f^c}\}^{-1} (Q \circ \mathcal{S})_{\mathcal{X}_f^c \mathcal{X}_f})^{-1}.$$

Hence for $x, y \in \mathcal{X}_f$, it follows by a similiar calculation leading to (8) that for arbitrary $d_u \geq 0,\ u \in \mathcal{X}_f$,

$$\mathbf{E}\left[Z_x^{\mathcal{S}} Z_y^{\mathcal{S}} e^{-\sum_{u \in \mathcal{X}_f} d_u \frac{(Z_u^{\mathcal{S}})^2}{2}} \right] = \mathbf{E}\left[e^{-\sum_{u \in \mathcal{X}_f} d_u \frac{(Z_u^{\mathcal{S}})^2}{2}} \right] (D_f + \Sigma_f^{-1})^{-1}(x, y).$$

Here D_f is a diagonal matrix of dimension $|\mathcal{X}_f|$ with diagonal entries $d_u,\ u \in \mathcal{X}_f$. We now prove a claim which will help us prove that $(D_f + \Sigma_f^{-1})^{-1}(x, y)$ is indeed $(D - Q \circ \mathcal{S})^{-1}(x, y)$, where D is a diagonal matrix of dimension $|\mathcal{X}|$ with diagonal entry d_u if $u \in \mathcal{X}_f$ and 0 otherwise.

Claim 4.1 *Let $A \subset \mathcal{X}$ be finite and*

$$(Q \circ \mathcal{S})^A := (Q \circ \mathcal{S})_{AA} - (Q \circ \mathcal{S})_{AB} \{(Q \circ \mathcal{S})_{BB}\}^{-1} (Q \circ \mathcal{S})_{BA}.$$

If Q satisfies (1) or (2),

$$(a) \qquad (Q \circ \mathcal{S})^A(a, a') = \begin{cases} q_{aa}(1 - \mathbf{E}_a[1_{\{X_{R_A} = a\}} H_{R_A}]) & \text{if } a = a' \\ -q_{aa} \mathbf{E}_a[1_{\{X_{R_A} = a'\}} H_{R_A}] & \text{if } a \neq a' \end{cases}$$

and,

$$(b) \qquad \{(Q \circ \mathcal{S})^A\}^{-1}(a, a') = (Q \circ \mathcal{S})^{-1}(a, a')\ \forall a, a' \in A.$$

Proof Throughout the proof, the absolute convergence for various infinite sums will be taken care of by the fact that $\sum_{n=0}^{\infty} P^n < \infty$ (because Q satisifes (1) or (2)). Note that the prediction formulas derived for the finite case in Section 4.1 go through for the infinite case as well, if A is finite. Recall that R_A is the first time the process $\{X_t\}_{t\geq 0}$ hits A after the initial state. Hence from (13) and strong Markov property,

$$\begin{aligned}
-(Q \circ \mathcal{S})^A(a, a') &= -q_{aa'} - \sum_{b\in B} q_{ab}\mathcal{S}(a, b)\mathbf{E}_b[1_{\{X_{R_A}=a'\}}H_{R_A}] \\
&= -q_{aa'} + q_{aa}\mathbf{E}_a[1_{\{X_{R_A}=a'\}}1_{\{R_A>S_1\}}H_{R_A}] \\
&= \begin{cases} -q_{aa}(1 - \mathbf{E}_a[1_{\{X_{R_A}=a\}}H_{R_A}]) & \text{if } a = a' \\ q_{aa}\mathbf{E}_a[1_{\{X_{R_A}=a'\}}H_{R_A}] & \text{if } a \neq a' \end{cases}
\end{aligned}$$

This completes the proof of (a). This also gives $(Q \circ \mathcal{S})^A = Q^{diag}_{AA}(I - P^A)$ where $P^A(a, a') := \mathbf{E}_a[1_{\{X_{R_A}=a'\}}H_{R_A}],\ \forall a, a' \in A$, and Q^{diag} is the diagonal matrix with the same diagonal entries as Q. Let R^n_A denote the time of n^{th} return to A. It follows that,

$$\begin{aligned}
\{(Q \circ \mathcal{S})^A\}^{-1}(a, a') &= \frac{1}{q_{aa}} \sum_{n=0}^{\infty} (P^A)^n(a, a') \\
&= \frac{1}{q_{aa}} \sum_{n=0}^{\infty} \sum_{\substack{a_0,a_1,\ldots,a_n\in A \\ a_0=a,a_n=a'}} \prod_{i=0}^{n-1} \mathbf{E}_{a_i}[1_{\{X_{R_A}=a_{i+1}\}}H_{R_A}] \\
&= \frac{1}{q_{aa}} \sum_{n=0}^{\infty} \mathbf{E}_a[1_{\{X_{R^n_A}=a'\}}H_{R^n_A}]
\end{aligned}$$

The previous equality follows by the definition of $\{H_t\}_{t\geq 0}$ and repeated application of the strong Markov property. Observing that $X_{S_i} = a'$ only if $S_i = R^n_A$ for some $n \geq 1$, we get that,

$$\begin{aligned}
\{(Q \circ \mathcal{S})^A\}^{-1}(a, a') &= \frac{1}{q_{aa}} \sum_{i=0}^{\infty} \mathbf{E}_a[1_{\{X_{S_i}=a'\}}H_{S_i}] \\
&= -\mathbf{E}_a\left[\int_0^{\infty} 1_{\{X_s=a'\}}H_s ds\right] \\
&= (Q \circ \mathcal{S})^{-1}(a, a')
\end{aligned}$$

The previous equality follows from Lemma 4.1. The proof of (b) is now complete. ■

Note that $-(D-Q)$ is a generator matrix that satisfies (1) or (2) (because Q satisfies one of these conditions). Applying Claim 4.1 for $-(D-Q)$ with $A = \mathcal{X}_f$, we get that,

$$(D_f + \Sigma_f^{-1})^{-1}(x,y) = ((D-Q)\circ\mathcal{S})^{-1}(x,y).$$

By imitating the proof of Lemma 4.2 we get

$$((D-Q)\circ\mathcal{S})^{-1}(x,y) = \int e^{-\sum_{u\in\mathcal{X}_f} d_u l_\infty^u} H_\infty d\mu_{xy}.$$

Combining everything,

$$\mathbf{E}\left[Z_x^{\mathcal{S}} Z_y^{\mathcal{S}} e^{-\sum_{u\in\mathcal{X}_f} d_u \frac{(Z_u^{\mathcal{S}})^2}{2}}\right] = \int \mathbf{E}\left[e^{-\sum_{u\in\mathcal{X}_f} d_u\left(\frac{(Z_u^{\mathcal{S}})^2}{2}+l_\infty^u\right)}\right] d\mu_{xy}.$$

Note that the set of functions $F_{\underline{d}}(\underline{w}) = e^{-\sum_{u\in\mathcal{X}} d_u w_u}$, where $\underline{d} = \{d_u\}_{u\in\mathcal{X}}$ is arbitrary with $d_u > 0$ for finitely many $u \in \mathcal{X}$ and $d_u = 0$ otherwise, generate the Borel σ-algebra in $\mathbb{R}^{|\mathcal{X}|}$ and they form a closed class under multiplication. Also, the set of functions F for which

$$\mathbf{E}\left[Z_x^{\mathcal{S}} Z_y^{\mathcal{S}} F\left(\frac{(Z_u^{\mathcal{S}})^2}{2},\ u\in\mathcal{X}\right)\right] = \int \mathbf{E}\left[F\left(\frac{(Z_u^{\mathcal{S}})^2}{2} + l_\infty^u,\ u\in\mathcal{X}\right) H_\infty\right] d\mu_{xy},$$

is a linear space closed under bounded convergence and under monotone convergence. Hence for each bounded Borel measurable function $F : \mathbb{R}^{|\mathcal{X}|} \to \mathbb{R}$,

$$\mathbf{E}\left[Z_x^{\mathcal{S}} Z_y^{\mathcal{S}} F\left(\frac{(Z_u^{\mathcal{S}})^2}{2},\ u\in\mathcal{X}\right)\right] = \int \mathbf{E}\left[F\left(\frac{(Z_u^{\mathcal{S}})^2}{2} + l_\infty^u,\ u\in\mathcal{X}\right) H_\infty\right] d\mu_{xy}.$$

4.3 Conditional Independence Property

There is another interesting property of Dynkin's isomorphism which is preserved after introducing a "sign" matrix $\mathcal{S}$. Let Q be the generator of a continuous time Markov process $\{X_t\}_{t\geq 0}$, with a countable state space $\mathcal{X}$. Assume Q^{-1} exists and Q is symmetric. Let $\{Z_x^{\mathcal{S}}\}_{x\in\mathcal{X}}$ be a zero mean Gaussian field with variance-covariance matrix $\Sigma^{\mathcal{S}} = (-Q\circ\mathcal{S})^{-1}$.

Lemma 4.3 *Let A, B, C be disjoint subsets of the state space $\mathcal{X}$, such that to go from any state in A to any state in C, the Markov process $\{X_t\}_{t\geq 0}$ has to pass through B. Then conditioned on $Z_B^{\mathcal{S}} := \{Z_b^{\mathcal{S}}\}_{b\in B}$, the Gaussian random vectors $Z_A^{\mathcal{S}} := \{Z_a^{\mathcal{S}}\}_{a\in A}$ and $Z_C^{\mathcal{S}} := \{Z_c^{\mathcal{S}}\}_{c\in C}$ are independent.*

Proof Fix $a \in A$ and $c \in C$ arbitrarily. By (15),

$$Cov(Z_a^{\mathcal{S}}, Z_c^{\mathcal{S}} \mid Z_b^{\mathcal{S}},\ b \in B) = \mathbf{E}_a \left[\int_0^\infty 1_{\{X_s = c\}} 1_{\{s < R_B\}} H_s ds \right].$$

Since the Markov process $\{X_t\}_{t \geq 0}$ has to pass through the set B to go from the state a to the state c,

$$1_{\{X_s = c\}} 1_{\{s < R_B\}} = 0 \text{ under } \mathbf{P}_a.$$

Hence,

$$Cov(Z_a^{\mathcal{S}}, Z_c^{\mathcal{S}} \mid Z_b^{\mathcal{S}},\ b \in B) = 0.$$

Since $a \in A$ and $c \in C$ were arbitrarily fixed, it follows that $Z_A^{\mathcal{S}}$ and $Z_C^{\mathcal{S}}$ are uncorrelated given $Z_B^{\mathcal{S}}$. Two random vectors having a joint Gaussian distribution are independent iff they are uncorrelated. Hence, $Z_A^{\mathcal{S}}$ and $Z_C^{\mathcal{S}}$ are independent given $Z_B^{\mathcal{S}}$. ∎

Bolthausen [1] uses this property in his analysis of the Gaussian free field.

4.4 An Example: Ornstein-Uhlenbeck Process on $\mathbb{N}$

Consider the Ornstein-Uhlenbeck process $\{Z_i\}_{i \in \mathbb{N}}$ defined by

$$Z_1 = \varepsilon_1,\ \ Z_i = aZ_{i-1} + \varepsilon_i\ \forall i \geq 2, \text{ where } \{\varepsilon_i\}_{i \geq 1} \text{ are i.i.d. } N(0,1).$$

Let Σ denote the variance-covariance matrix of $\{Z_i\}_{i \in \mathbb{N}}$. Then,

$$\Sigma(k,l) = \begin{cases} a^{l-k} \sum_{i=1}^{k} a^{2(k-i)} & \text{if } k \leq l, \\ \Sigma(l,k) & \text{if } l < k. \end{cases} \tag{16}$$

After some manipulations, we can establish that $\Sigma = -Q^{-1}$ where

$$Q(k,l) = \begin{cases} -(1+a^2) & \text{if } k = l, \\ a & \text{if } k = l \pm 1, \\ 0 & \text{otherwise.} \end{cases} \tag{17}$$

If $a > 0$, then Q is the generator of a birth and death process. Hence, the Ornstein-Uhlenbeck process is connected to the birth and death process with generator Q by Dynkin's isomorphism.

Suppose we introduce a "sign" matrix $\mathcal{S}$ (as described earlier in this section) and for $a > 0$, look at a Gaussian field $\{Z'_i\}_{i\in\mathbb{N}}$ defined by

$$Z'_1 = \varepsilon_1, \;\; Z'_i = \mathcal{S}(i-1,i)aZ'_{i-1} + \varepsilon_i, \; \forall i \geq 2, \text{ where } \{\varepsilon_i\}_{i\geq 1} \text{ are i.i.d. } N(0,1).$$

It follows after some manipulations that the variance-covariance matrix of $\{Z'_i\}_{i\in\mathbb{N}}$ is given by $(-Q \circ \mathcal{S})^{-1}$, where Q is as specified in (17).
If $\mathcal{S}(k,l) = -1 \; \forall k \neq l$, then

$$Z'_1 = \varepsilon_1, \;\; Z'_i = -aZ'_{i-1} + \varepsilon_i, \; \forall i \geq 2.$$

Hence, if $a > 0$, then the Ornstein-Uhlenbeck process on $\mathbb{N}$ with parameter $-a$ is associated to a birth and death process with generator Q in (17), by Dynkin's isomorphism with "sign" matrix $\mathcal{S}$ such that $\mathcal{S}(k,l) = -1 \; \forall k \neq l$.

5 An Algorithm for Computing the Prediction Coefficients

We present an algorithm for computing the prediction coefficients $\mathbf{E}_b\left[1_{\{X_{R_A}=a\}}H_{R_A}\right]$ in (14) for calculating $\mathbf{E}[Z^{\mathcal{S}}_b \mid Z^{\mathcal{S}}_a, \; a \in A]$. This algorithm can be described in two ways:

- Graph theoretic description.

  Construct a graph on the vertex set $\mathcal{X}$ by putting an edge of weight $Q(x,y)\mathcal{S}(x,y)$ between vertices x and y, $\forall x \neq y \in \mathcal{X}$. If any of these weights are 0, that by default means no edge is put between the corresponding vertices. Put a loop of weight $-Q(x,x)$ at each vertex $x \in \mathcal{X}$. Since we want to predict $Z^{\mathcal{S}}_b$ given $\{Z^{\mathcal{S}}_a\}_{a\in A}$ we now proceed to remove all vertices not in $\{b\} \cup A$ from this graph in a sequential fashion. Choose any vertex, say z not in $\{b\} \cup A$. If we remove z, i.e. if we behave as if z does not exist in the graph, this leads to forming a new edge between every pair x and y such that x and z, as well as y and z share an edge. The weight of this new edge is the product of the weights of these two edges divided by the weight of the loop at z. If there is already an edge between x and y, add the weight of this new edge to the existing one and combine them into one edge. Perform this procedure with all x and y sharing an edge with z (including the case

$x = y$). So we get a new graph with vertex set $\mathcal{X} \setminus \{z\}$ and edge set as described above. Note that $-Q \circ \mathcal{S}$ is a diagonally dominant matrix with positive diagonal entries and hence for the old graph, the weight of the loop at any vertex dominates the sum of the absolute weight values of the the edges emanating from that vertex. As we will see later, the new grpah has the same property. We continue choosing vertices and removing them by using the above procedure until we are left with the vertex set $A \cup \{b\}$. The coefficient of $Z_a^{\mathcal{S}}$ in $\mathbf{E}[Z_b^{\mathcal{S}} \mid Z_a^{\mathcal{S}},\ a \in A]$ is precisely the weight of the edge joining a and b divided by the weight of the loop at b, for every a in A.

- Analytic description.

  We can describe the above algorithm analytically as follows:

  1. Start with $V = \mathcal{X},\ M = -Q \circ \mathcal{S}$.
  2. Choose $z \in V \setminus \{A \cup \{b\}\}$.
  3. $M(x,y) = M(x,y) - \frac{M(x,z)M(y,z)}{M(z,z)}\ \forall x, y \in V$.
  4. Remove the z^{th} row and the z^{th} column of M.
  5. $V \to V \setminus \{z\}$. If $V \neq A \cup \{b\}$ goto step 2, otherwise stop. The coefficient of $Z_a^{\mathcal{S}}$ in $\mathbf{E}[Z_b^{\mathcal{S}} \mid Z_a^{\mathcal{S}},\ a \in A]$ is $-\frac{M(a,b)}{M(b,b)}$ for every $a \in A$.

The above description tells us that our algorithm is essentially the sequential process of evaluating the Schur complement $(-Q \circ \mathcal{S})_{VV} - (-Q \circ \mathcal{S})_{VV^c}\{(-Q \circ \mathcal{S})_{V^cV^c}\}^{-1}(-Q \circ \mathcal{S})_{V^cV}$ (and ending at $V = A \cup \{b\}$) by reducing rows and columns. Since the Schur complement of a diagonally dominant matrix is also diagonally dominant, the matrix M is a diagonally dominant matrix at every step of the algorithm. The proof of this algorithm can be obtained immediately by observing two facts. Firstly, $\Sigma^{\mathcal{S}} = (-Q \circ \mathcal{S})^{-1}$ implies that

$$(\Sigma_{VV}^{\mathcal{S}})^{-1} = (-Q \circ \mathcal{S})_{VV} - (-Q \circ \mathcal{S})_{VV^c}\{(-Q \circ \mathcal{S})_{V^cV^c}\}^{-1}(-Q \circ \mathcal{S})_{V^cV} \text{ for every } V \subseteq \mathcal{X}.$$

Hence, when we stop the algorithm, the matrix M is the same as $(\Sigma_{VV}^{\mathcal{S}})^{-1}$ with $V = A \cup \{b\}$. Secondly, if $\underline{Y} \sim MVN_n(\underline{0}, \Gamma)$, then

$$\mathbf{E}[Y_i \mid Y_j,\ j \neq i] = \sum_{j \neq i} \frac{-\Gamma^{-1}(i,j)}{\Gamma^{-1}(i,i)} Y_j. \tag{18}$$

Since $\{Z_v^{\mathcal{S}}\}_{v\in V}$ is $MVN_{|V|}(\underline{0}, \Sigma_{VV}^{\mathcal{S}})$ and $M = (\Sigma_{VV}^{\mathcal{S}})^{-1}$, it follows by (18) that

$$\mathbf{E}[Z_b^{\mathcal{S}} \mid Z_a^{\mathcal{S}},\ a \in V \setminus \{b\}] = \sum_{a\in V\setminus\{b\}} \frac{-M(a,b)}{M(b,b)} Z_a^{\mathcal{S}}.$$

Hence this algorithm is not all that mysterious. If $|\mathcal{X}| = n$, the worst case running time of this algorithm is $O(n^3)$. One nice property of this algorithm is that at any step of the algorithm with vertex set V and corresponding matrix M,

$$\mathbf{E}[Z_v^{\mathcal{S}} \mid Z_w^{\mathcal{S}},\ w \in V \setminus \{v\}] = \sum_{w\in V\setminus\{v\}} \frac{-M(v,w)}{M(v,v)} Z_w^{\mathcal{S}}.$$

Hence, if we want we can obtain the prediction coefficients given $\{Z_v^{\mathcal{S}}\}_{v\in V}$ for every $V \subseteq \mathcal{X}$ that comes up in the course of this algorithm.

5.1 An Example of Prediction with Independent Errors at the Observed Values

Consider the Ornstein-Uhlenbeck process on $\mathbb{N}$ with $a = 1$, i.e.

$$Z_1 = \varepsilon_1,\ \ Z_i = Z_{i-1} + \varepsilon_i\ \forall i \geq 2,\ \text{where } \{\varepsilon_i\}_{i\geq 1} \text{ are i.i.d. } N(0,1).$$

This process is same as the Gaussian free field on $\mathbb{N}$. It follows that the variance-covariance matrix Σ of $\{Z_i\}_{i\in\mathbb{N}}$ is given by

$$\Sigma(k,l) = k \wedge l\ \forall k,l \in \mathbb{N}.$$

Suppose we observe the values of the process in the set $V = \{n_1, n_2, ..., n_k\}$ (where $n_i < n_j$ if $i < j$), but with an independent additive error $\tilde{\varepsilon}_i$ at the point n_i $i = 1,2,...,k$, where $\{\tilde{\varepsilon}_i\}_{1\leq i\leq k}$ are i.i.d. $N(0,\sigma^2)$. With these observations, we want to predict the process $\{Z_i\}_{i\in\mathbb{N}}$ i.e. we want to compute the expectation

$$\mathbf{E}[Z_n \mid Z_{n_1} + \tilde{\varepsilon}_1, Z_{n_2} + \tilde{\varepsilon}_2, ..., Z_{n_k} + \tilde{\varepsilon}_k],\ \forall n \in \mathbb{N}.$$

It is known that

$$\mathbf{E}[Z_n \mid Z_{n_i} + \tilde{\varepsilon}_i,\ i = 1,2,...,k] = \Sigma_{nV}(\Sigma_{VV} + \sigma^2 I_{|V|})^{-1} Z_V.$$

Since $\Sigma(n_i, n) = n_i \wedge n \ \forall i = 1, 2, ..., k$, we would like to compute a simplified expression for $(\Sigma_{VV} + \sigma^2 I_{|V|})^{-1}$. We utilize the structure of Σ_{VV} for this purpose.

$$\Sigma_{VV} = UDU^T,$$

where, $U(i,j) = 1_{\{i \geq j\}}, \ \forall 1 \leq i,j \leq k$ and D is a diagonal matrix with $D(i,i) = n_i - n_{i-1} \ \forall 1 \leq i \leq k$ (where $n_0 = 0$). It follows that

$$\Sigma_{VV} + \sigma^2 I_{|V|} = U \Lambda U^T,$$

where Λ is the tridiagonal matrix with

$$\Lambda(i,j) = \begin{cases} n_i - n_{i-1} + i\sigma^2 & \text{if } i = j, \\ -\sigma^2 & \text{if } |i-j| = 1, \\ 0 & \text{otherwise.} \end{cases} \tag{19}$$

Note that,

$$U^{-1}(i,j) = \begin{cases} 1 & \text{if } i = j, \\ -1 & \text{if } i = j+1, \\ 0 & \text{otherwise.} \end{cases} \tag{20}$$

Let $r_0 = \sigma^2, \ r_1 = (n_1 + \sigma^2)\sigma^2, \ r_i = (n_i - n_{i-1} + i\sigma^2) r_{i-1} - r_{i-2}$ for $i = 2, 3, ..., k$. By the explicit formula for the inverse of a symmetric tridiagonal matrix in [12],

$$\Lambda^{-1}(i,j) = \begin{cases} \frac{r_{i-1} r_{k-j}}{r_k} & \text{if } i \leq j, \\ \Lambda^{-1}(j,i) & \text{if } i > j. \end{cases} \tag{21}$$

Hence we obtain

$$\mathbf{E}[Z_n \mid Z_{n_i} + \tilde{\varepsilon}_i, \ i = 1, 2, ..., k] = \sum_{i=1}^{k} (\sum_{j=1}^{k} \gamma_{ij} (n \wedge n_j))(Z_{n_i} + \tilde{\varepsilon}_i),$$

where,

$$\gamma_{ij} = \begin{cases} \frac{(r_i - r_{i-1})(r_{k-j} - r_{k-j-1})}{r_k} & \text{if } i < j, \\ \frac{r_{i-1}(r_{k-i} - 2r_{k-i-1}) + r_i r_{k-i-1}}{r_k} & \text{if } i = j, \\ \gamma_{ji} & \text{if } i > j. \end{cases} \tag{22}$$

As is clear from this example, introducing errors leads to non-trivial changes in the prediction coefficients. It is hard to find a general formula which expresses these changed coefficients in terms of the associated Markov chain.

References

[1] E. Bolthausen (2000). Random walk representations and entropic repulsion for gradient models, In *"Infinite Dimensional Analysis"*, Royal Netherlands academy of Arts and Sciences, Amstredam, 55-84.

[2] Diaconis, P. and Evans, S. (2002). A different construction of Gaussian fields from Markov chains: Dirichlet Covariances, *Ann. Inst. Henri Poincare* **38**, 863-878.

[3] Durrett, R. (1996). *Probability: thoery and examples (*2^{nd} *edition)*, Wadsworth Publisihing Company, Belmont, CA.

[4] E.B. Dynkin (1980). Markov processes and random fields, *Bull. Amer. Math. Soc. (N.S.)* **3**, 975-999.

[5] E.B. Dynkin (1983). Markov processes as a tool in field theory, *J. Funct. Anal.* **50**, 167-187.

[6] E.B. Dynkin (1984). Gaussian and non-Gaussian random fields associated with Markov processes, *J. Funct. Anal.* **55**, 344-376.

[7] E.B. Dynkin (1984). Polynomials of the occupation field and related random fields, *J. Funct. Anal.* **58**, 20-52.

[8] N. Eisenbaum (1994). Dynkin's isomorphism and the Ray-Knight theorems, *Probability Theory and Related Fields* **99**, 321-335.

[9] N. Eisenbaum (1995). Une version sans conditionnement du thèoreme d'isomorphisms de Dynkin, In *"Seminaire de Probabilities, XXIX"*, Lecture Notes in Math., Vol. 1613, Springer, Berlin, 266-289.

[10] Eisenbaum, N., Kaspi, H., Marcus, M., Rosen, J. and Shi, Z. (2000). A Ray-Knight theorem for symmetric Markov processes, *Annals of Probability* **28**, 1781-1796.

[11] M.B. Marcus and J. Rosen (2006). *Markov Processes, Gaussian processes and Local Times*, Cambridge University Press.

[12] P. Schlegel (1970). The explicit inverse of a tridiagonal matrix, *Mathematics of Computation* **24**, 665-665.

[13] P. Sheppard (1985). On the Ray-Knight Markov property of local times, *J. London Math. Soc.* **31**, 377-384.

[14] D. Ylvisaker (1987). Prediction and design, *Ann. Statist.* **15**, 1-19.